Westfälische Wilhelms-Universität Münster

ERCIS
European Research Center for Information Systems

PREDICTIVE MODELING FOR BUSINESS PROCESSES

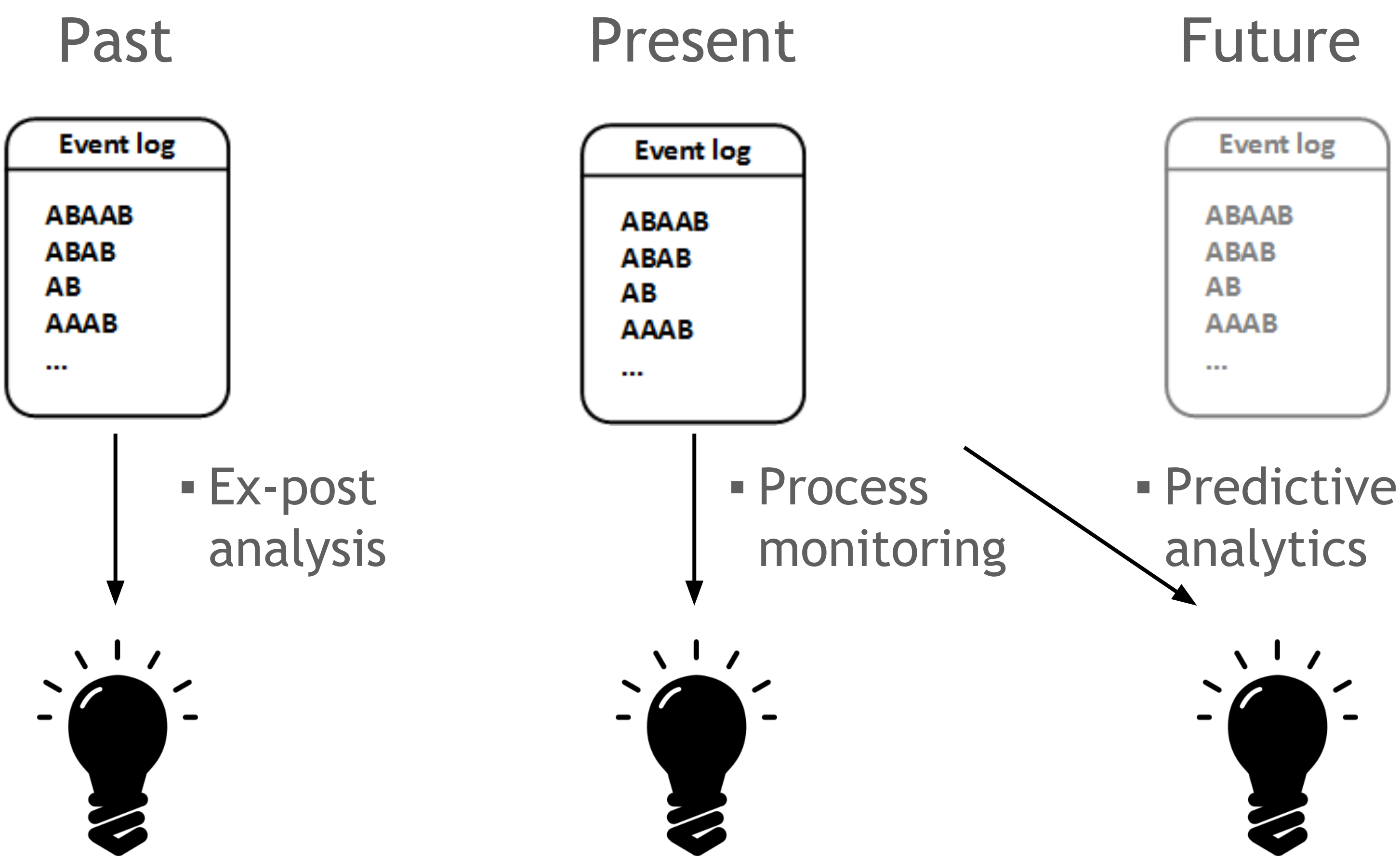
DESIGNING AND EVALUATING AN INTERPRETABLE PREDICTIVE MODELING TECHNIQUE FOR BUSINESS PROCESSES

DeMiMoP'14

Dominic Breuker

2014-09-08

MOTIVATION


Past
Present
Future
Event log
ABAAB
ABAB
AB
AAAB
...
Event log
ABAAB
ABAB
AB
AAAB
...
Event log
ABAAB
ABAB
AB
AAAB
...
▪ Ex-post analysis
▪ Process monitoring
▪ Predictive analytics

MOTIVATION

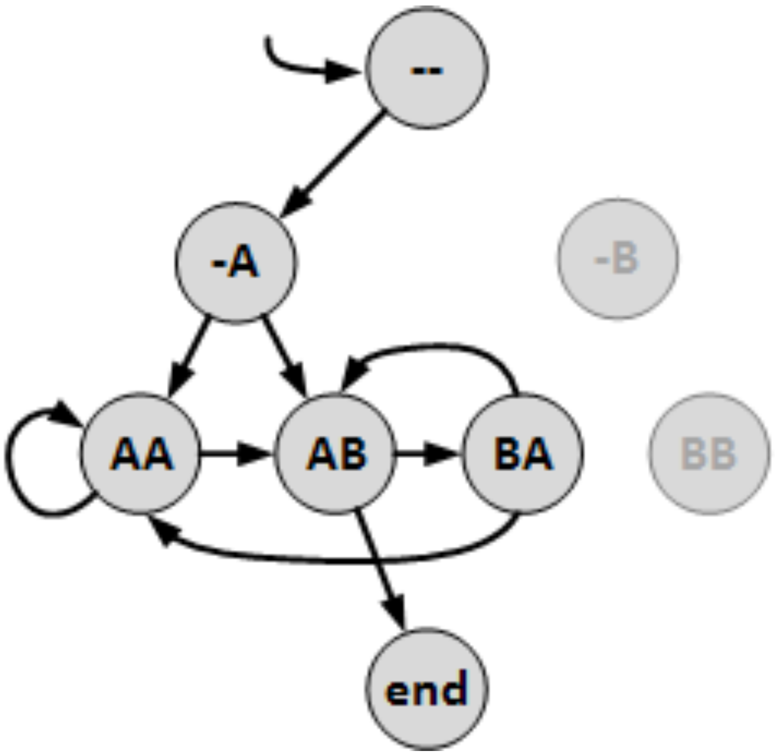
PREDICTIVE MODELING FOR BUSINESS PROCESSES? WHY?

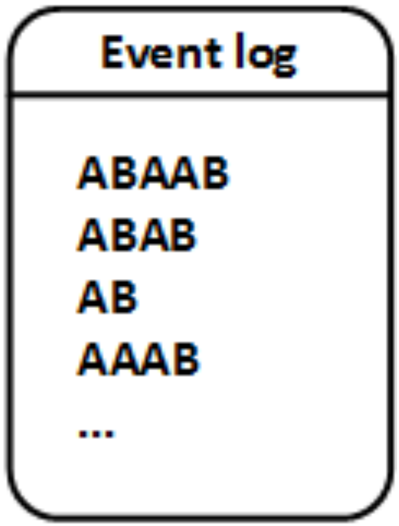
- Early warning systems
  - Predict future behavior
  - Warn managers if future is bad
  - Intervention possible

- Anomaly detection systems
  - Predict future behavior
  - Warn managers if a surprising future has happened
  - Analysis / Intervention possible

LITERATURE

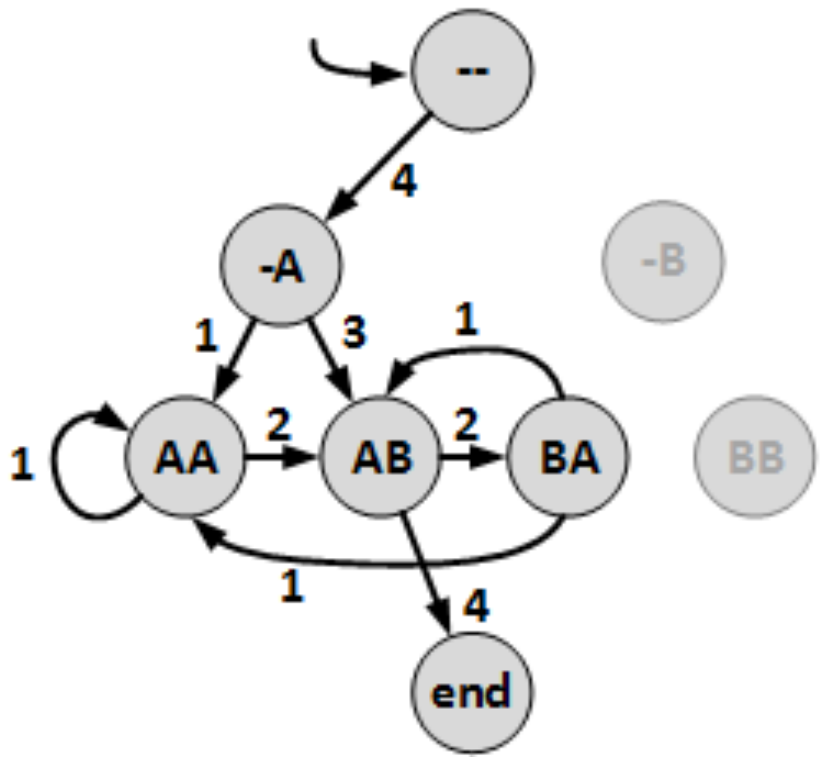
TRANSITION SYSTEM MINING





LITERATURE

FREQUENCIES ANNOTATED

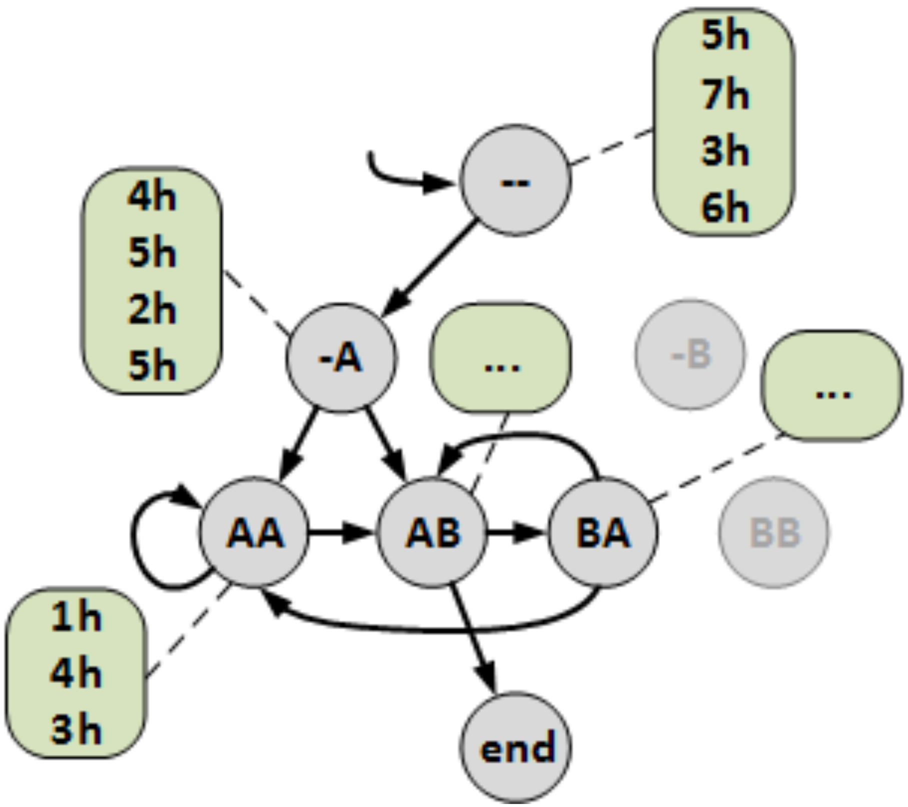


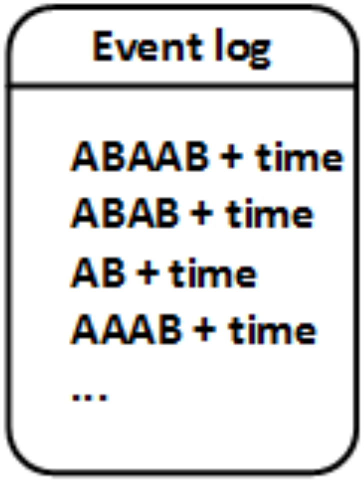
Event log

ABAAB
ABAB
AB
AAAB
...

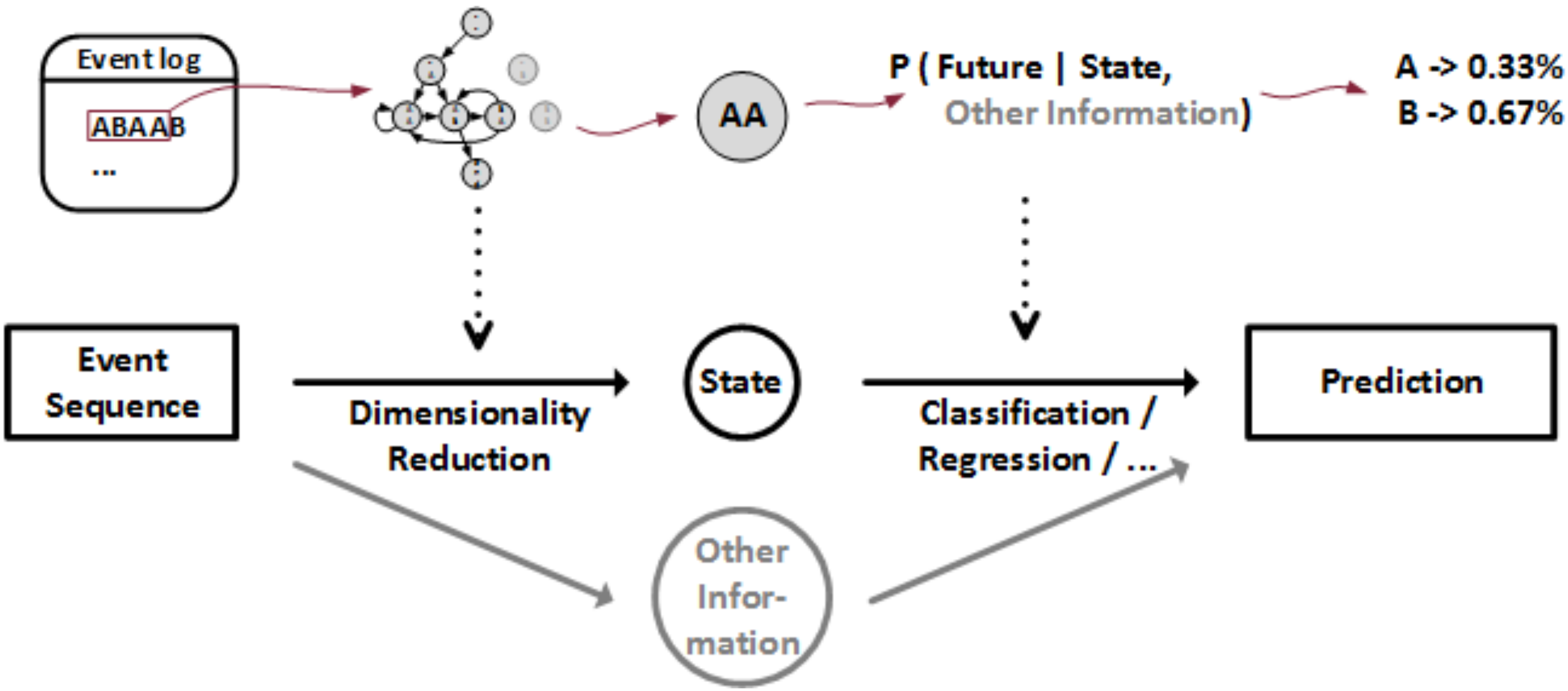
LITERATURE

COMPLETION TIMES ANNOTATED





ABSTRACT PRINCIPLE

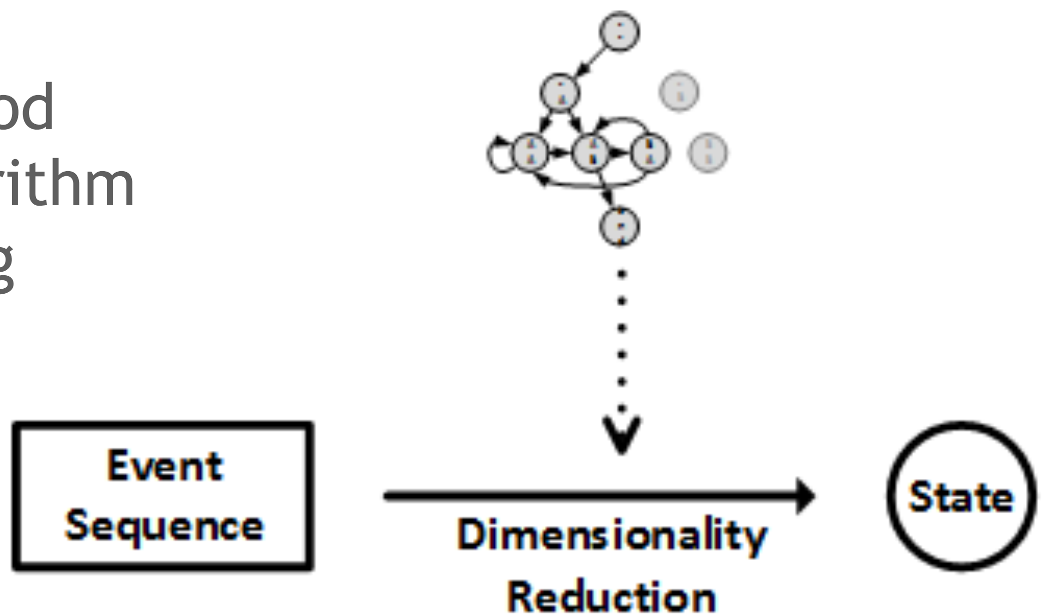


MAIN PROBLEM

DIMENSIONALITY REDUCTION <-> PROCESS DISCOVERY

- Dimensionality reduction
  - Map event log to a useful feature set

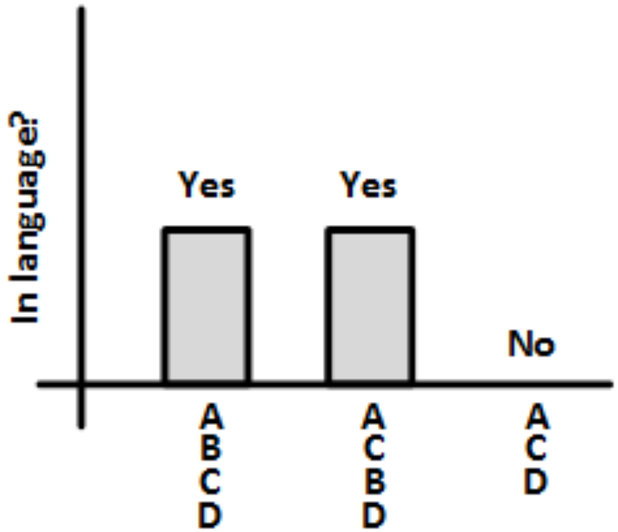
- Question: What is a good process discovery algorithm for predictive modeling applications?



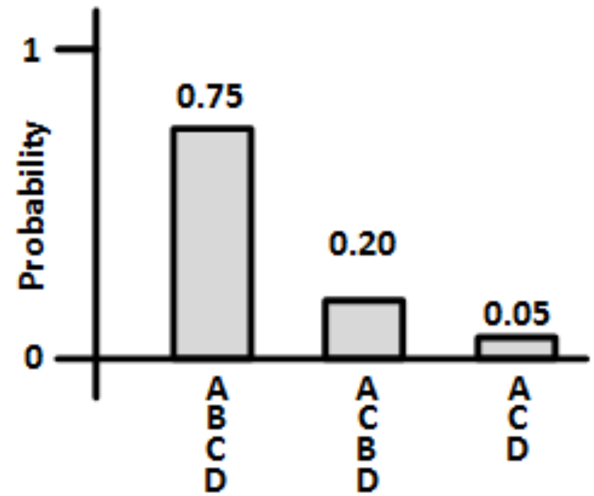
TWO APPROACHES

GRAMMATICAL INFERENCE THEORY

- Process = set of valid event sequences



- Process = probability distribution over event sequences

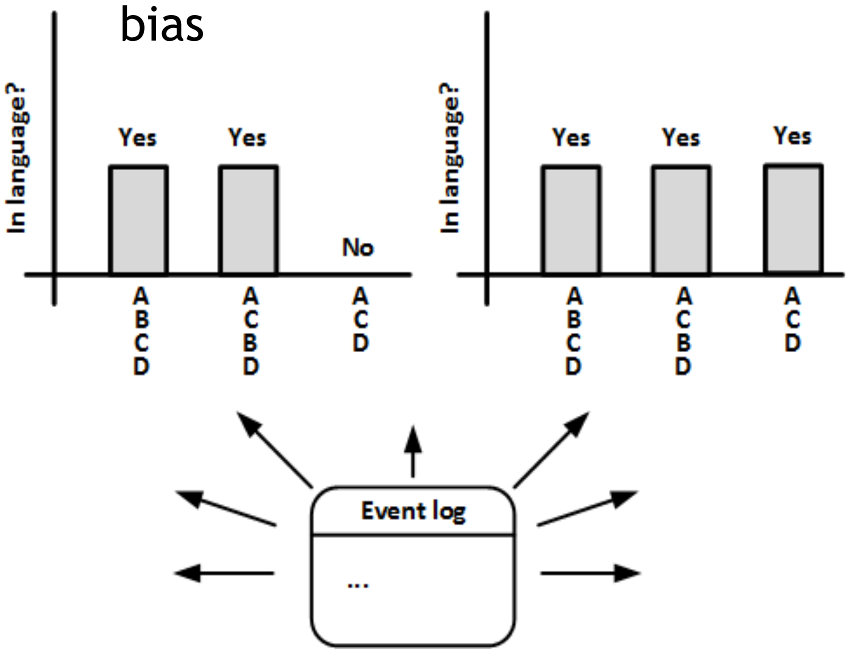


TWO APPROACHES

GRAMMATICAL INFERENCE THEORY

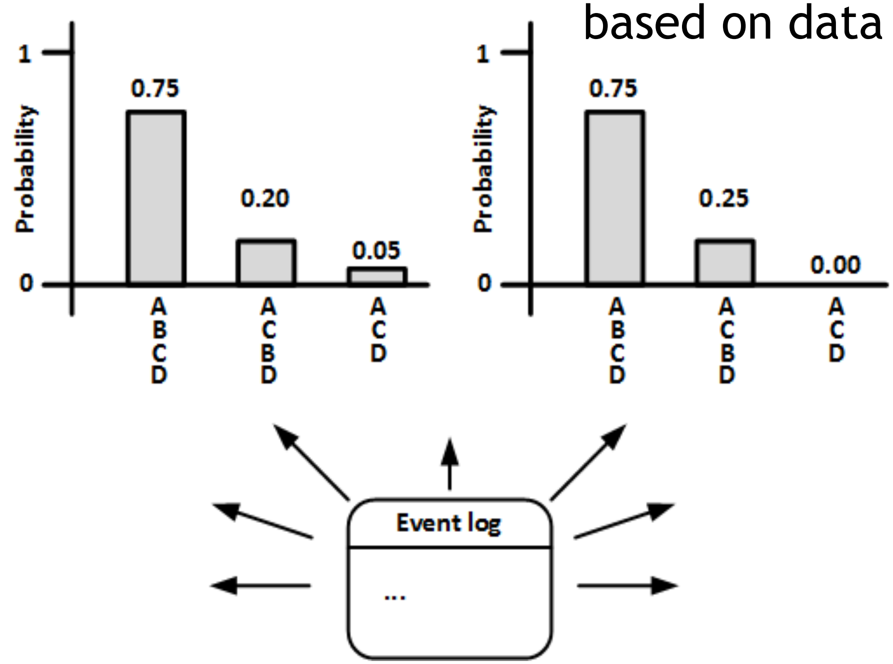
- Strong language bias necessary

Choice by language bias

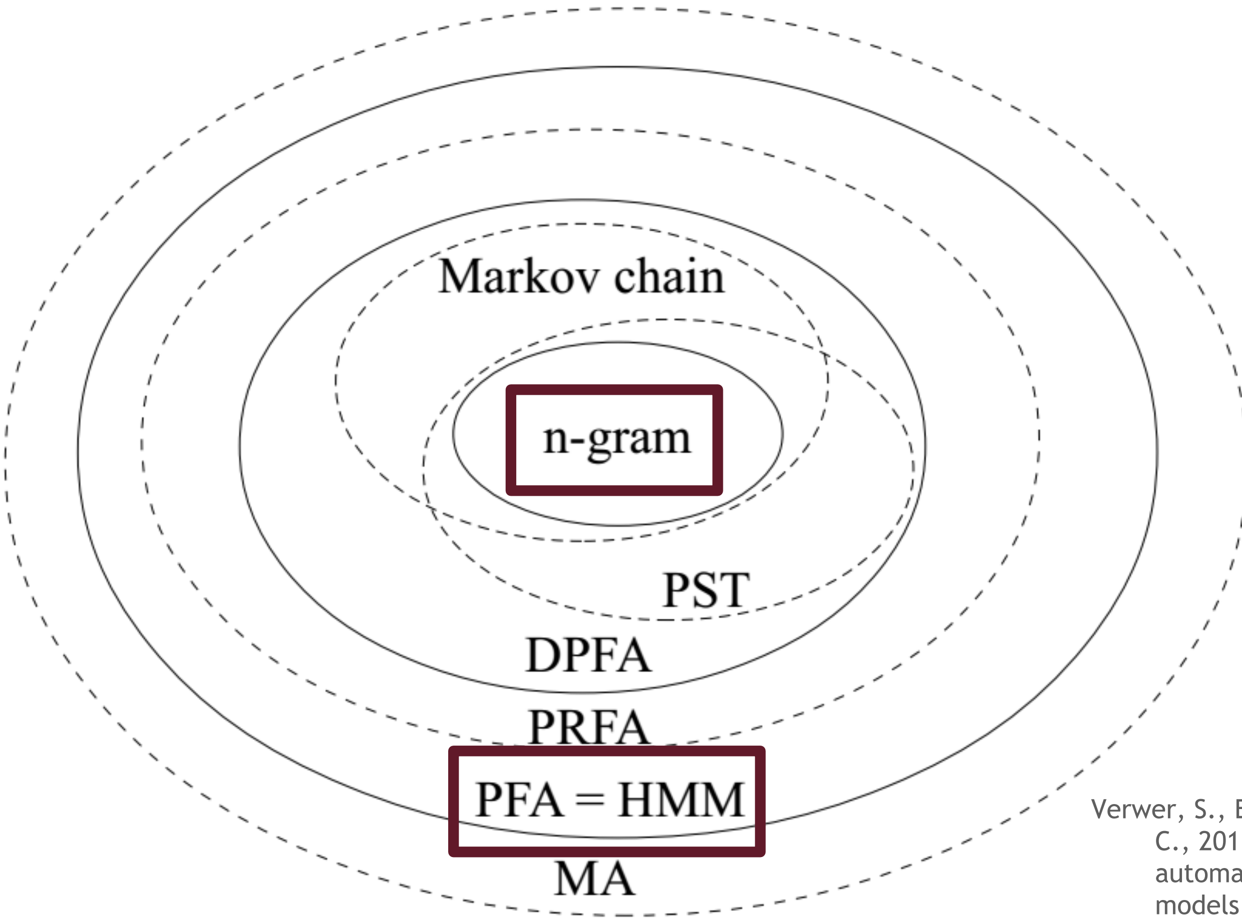


- Weaker language bias possible

Choice by comparison based on data



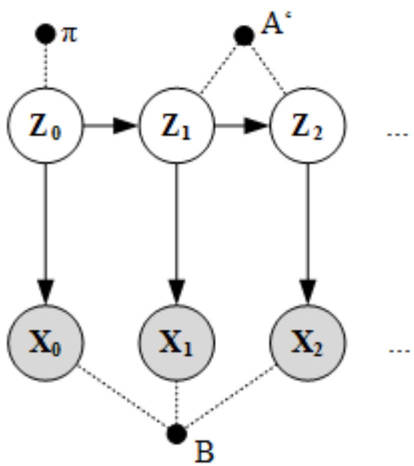
PROBABILISTIC MODELS



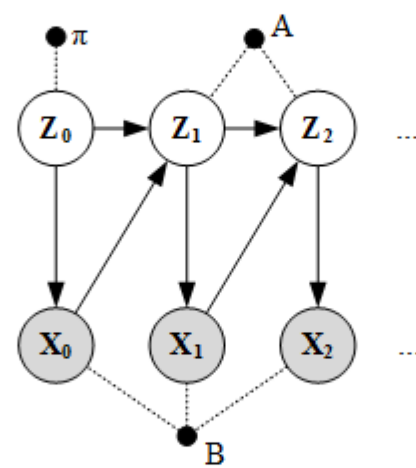
Verwer, S., Eyraud, R. & de la Higuera, C., 2013. PAutomaC: a probabilistic automata and hidden Markov models learning competition. *Machine Learning*.

PROBABILISTIC MODELS

- Hidden Markov Model (HMM)



- Probabilistic Finite Automaton (PFA)



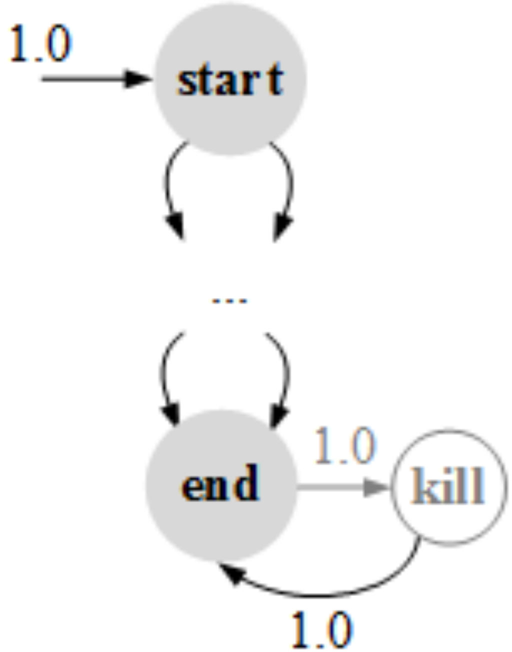
PFA ESTIMATION

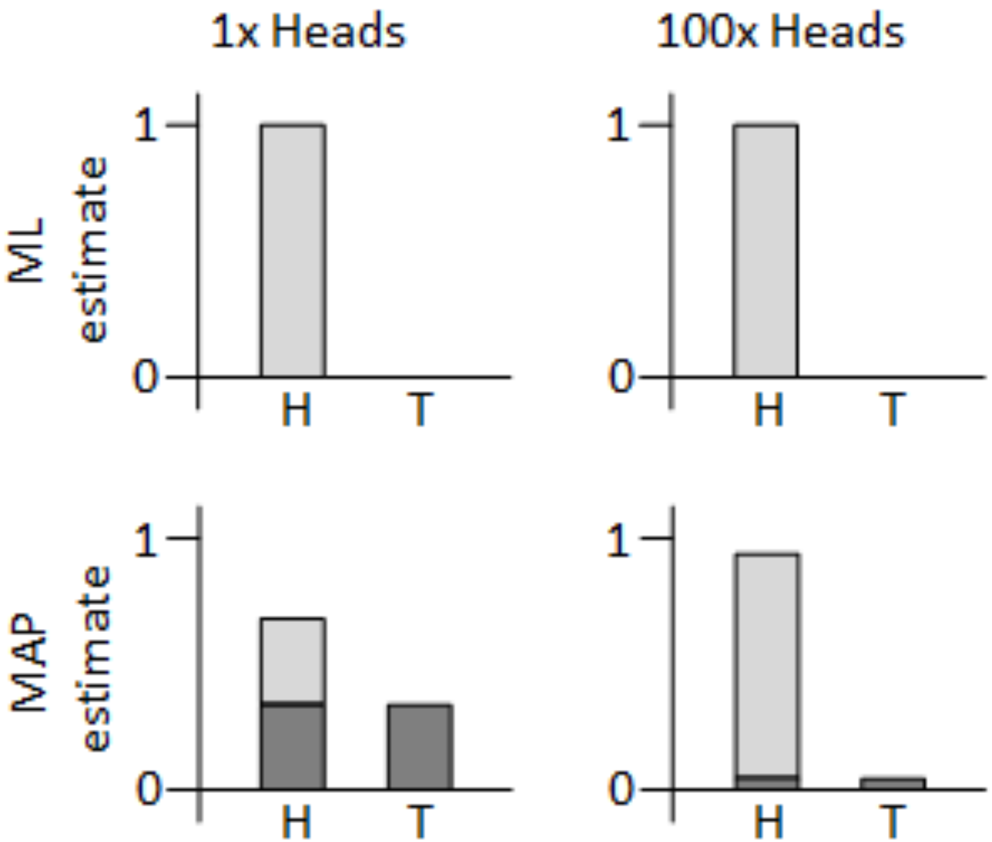
3 FAMILIES OF METHODS

- Bayesian inference
  - Do not estimate a single model (e.g., Gibbs sampling)
  - But: effective!
- Parameter estimation
  - Estimate parameters (often: ML estimation)
  - Quite effective too
- State merging
  - Iteratively merge states, starting with prefix tree
  - Least effective

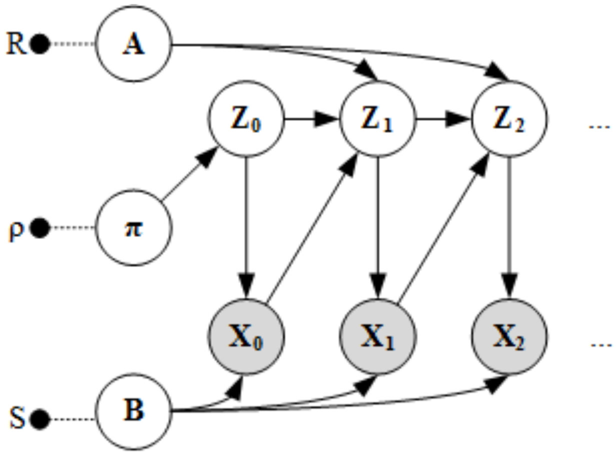
Verwer, S., Eyraud, R. & de la Higuera, C., 2013. PAutomaC: a probabilistic automata and hidden Markov models learning competition. *Machine Learning*.

PFA MODIFICATIONS





PFA MODIFICATIONS



$$P(Z_0) \sim \quad Categorical(\pi_0, \dots, \pi_K)$$

$$P(X_t | Z_t = k) \sim \quad Categorical(b_{k0}, \dots, b_{kE})$$

$$P(Z_t | Z_{t-1} = k, X_{t-1} = e) \sim \quad Categorical(a_{ke0}, \dots, a_{keK})$$

$$P(\pi_1, \dots, \pi_K) \sim \quad Dirichlet(\rho_1, \dots, \rho_K)$$

$$P(b_{k1}, \dots, b_{kE}) \sim \quad Dirichlet(s_{k1}, \dots, s_{kE})$$

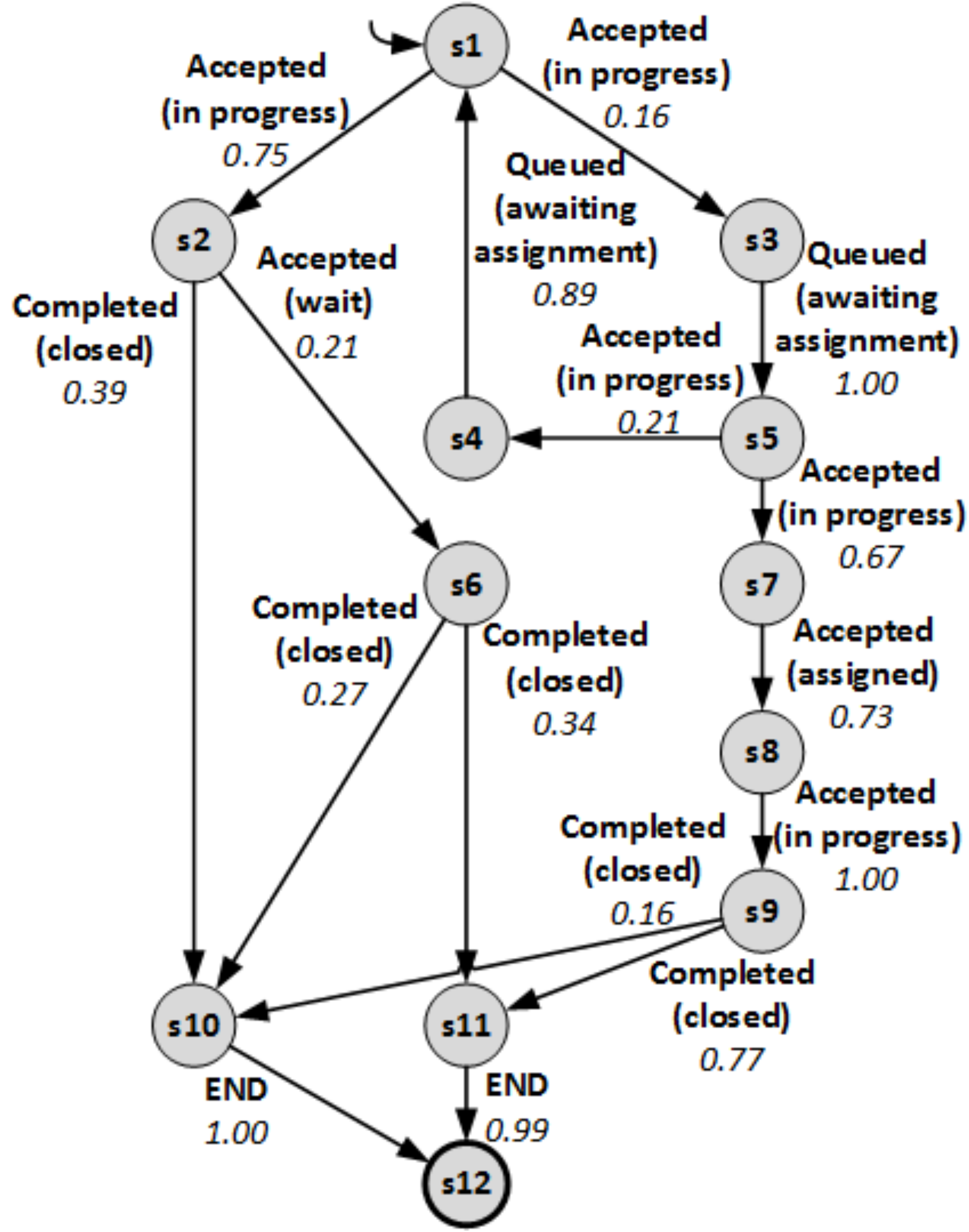
$$P(a_{ke1}, \dots, a_{keK}) \sim \quad Dirichlet(r_{ke1}, \dots, r_{keK})$$

EVALUATION (PREDICTION)

Event log	Predictor	Accuracy	ØSensitivity	ØSpecitivity	H
2012 W	EM	0.719	0.578	0.955	11.183
	5-gram	0.728	0.588	0.957	Infinity
2012 A	EM	0.801	0.723	0.980	3.093
	4-gram	0.801	0.723	0.980	2.839
2012 O	EM	0.811	0.647	0.973	4.513
	3-gram	0.811	0.647	0.973	4.180
2013 Incidents	EM	0.714	0.383	0.974	12.041
	4-gram	0.635	0.377	0.967	Infinity
2013 Problems	EM	0.690	0.521	0.945	7.231
	3-gram	0.699	0.564	0.948	Infinity

DEMONSTRATION

- Visualization is possible
- Threshold
  - Cut out improbable transitions
- Also possible: Petri net synthesis
  - -> Petrify



EVALUATION (PROCESS DISCOVERY)

Algorithm	Fitness	Advanced behavioral appropriateness
EM + Petrify	0.998	0.908
AGNES-Miner	0.995	0.813
α+	0.969	0.873
α++	0.984	0.879
DT Genetic Miner	0.996	0.778
Genetic Miner	0.998	0.737
HeuristicsMiner	0.973	0.809
ILP Miner	1.000	0.786

De Weerdt, J. et al., 2012. A multi-dimensional quality assessment of state of-the-art process discovery algorithms using real-life event logs. *Information Systems*, 37(7), pp.654–676.

Conclusion

- Goal: Devlelop a good “event sequence -> state” reduction for predictive modeling in BPM
  - Probabilistic approach
  - Weak language bias

- Probabilistic finite automaton (PFA)
  - Modified (start/end state + regularization)
  - Estimation with EM
  - Can be used as process discovery algorithm

- PFA can be better than n-gram approaches…
- … but does not have to be!

Questions?

Slides available at:

http://goo.gl/Bi99Ck